

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 23:56:58 ; Search time 3074 Seconds

(without alignments)  
10736.038 Million cell updates/sec

Title: US-10-091-628-1

Perfect score: 1134  
Sequence: 1 atgagagcattgtccag.....acatcattcatgcatag 1134

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

ched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pa:\*  
7: gb\_pl:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_scs:\*  
11: gb\_ey:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_da:\*  
15: em\_fun:\*  
16: em\_hum:\*  
17: em\_in:\*  
18: em\_mu:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pa:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_scs:\*  
27: em\_un:\*  
28: em\_vl:\*  
29: em\_hcg\_hum:\*  
30: em\_hcg\_inv:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_inv:\*  
33: em\_hcg\_inv:\*  
34: em\_hcg\_inv:\*  
35: em\_hcg\_inv:\*  
36: em\_hcg\_inv:\*  
37: em\_hcg\_inv:\*  
38: em\_hcg\_inv:\*  
39: em\_hcg\_inv:\*  
40: em\_hcg\_inv:\*  
41: em\_hcg\_inv:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	377	33.2	23618	9	AC079237 Homo sapi
C 2	377	33.2	192263	2	AC093827 Homo sapi
3	375.4	33.1	65266	2	AC099847 Homo sapi
4	320.4	28.3	2263	6	AR033870 Sequence
5	320.4	28.3	2263	6	AR033870 Sequence
6	320.4	28.3	2263	6	IG2744 Sequence 1
7	307.8	27.1	1629	10	U02028 Cricetus
8	307.8	27.1	1629	10	DB7059 House mouse
9	306.4	27.0	1116	4	AB002693 Mus muscu
10	297.8	26.3	1047	6	Z54357 O. cuniculus
11	297.8	26.3	1047	6	AR033871 Sequence
12	297.8	26.3	1047	6	132745 Sequence 3
13	293.2	25.9	4269	10	U10417 Homo sapien
14	265	23.4	106811	2	AC120684 Rattus norv
15	261.8	23.1	215210	10	AL713989 Mouse DNA
C 16	217	19.1	65958	9	AC105413 Homo sapi
17	184	16.2	1212	4	OCU131361 Oryctolag
18	183.2	16.2	1663	6	AX401950 Sequence
19	182.6	16.1	1663	10	RATSBACT
C 20	182.6	16.1	543	11	G51602 Rattus norv
21	173.8	15.3	1411	10	MMU95132 Mus musculu
22	173.8	15.3	1579	10	MMU95131 Mus musculu
23	173.8	15.3	1596	10	AB003303 Mouse mRN
24	173.6	15.3	1649	10	BC021154 Mus muscu
25	173.6	15.3	1580	6	AX409529 Sequence
26	173.6	15.3	1580	6	HUMANTEP
27	138.2	12.2	1714	9	BC012048 Homo sapi
28	138.2	12.2	1714	9	BC019066 Homo sapi
29	133.2	11.7	27247	4	OCU2005 Oryctolag
30	118.8	10.5	2020	9	HS18DBAL U67669 Human l1eal
31	118.8	10.5	7979	9	HSBARTPRO Z54350 Homo sapien
32	118.8	10.5	157352	2	AC009687 Homo sapi
C 33	118.8	10.5	170407	2	AC011177 Homo sapi
C 34	118.8	10.5	171132	2	AL161771 Human DNA
C 35	118.8	10.5	172835	2	AL139797 Homo sapi
36	117	10.3	1364	10	AB078635 Mus muscu
37	117	10.3	173363	2	AC114998 Mus muscu
C 38	117	10.3	182230	2	AC110744 Mus muscu
39	115.6	10.2	562	10	AF181258 Mesocricet
40	115.4	10.2	1790	9	BC004966 Homo sapi
41	115.4	10.2	4379	9	HSP3 X12458 Human P3 ge
C 42	115.4	10.2	137217	2	HSR211110 Homo sapi
C 43	115.4	10.2	160656	2	AL592072 Homo sapi
C 44	115.4	10.2	219447	9	HMFLNGEPD L44140 Homo sapien
45	113.8	10.0	1443	10	AF266724 Mus muscu

#### ALIGNMENTS

RESULT 1  
AC079237/23618 bp DNA linear PRI 21-FEB-2002  
LOCUS AC079237  
DEFINITION Homo sapiens BAC clone RP11-711J3 from 4, complete sequence.  
ACCESSION AC079237  
VERSION AC079237.7 GI:18482358  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 23618)  
AUTHORS Sulston, J.E. and Waterston, R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE 99063792  
 PUBMED 9847074  
 REFERENCE 2 (bases 1 to 23618)  
 AUTHORS Radionenko, M. and Meyer, R.  
 TITLE The sequence of Homo sapiens BAC clone RP11-711J3  
 JOURNAL Unpublished (2001)  
 REFERENCE 3 (bases 1 to 23618)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-AUG-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 23618)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 23618)  
 AUTHORS Waterston, R.  
 TITLE Direct Submission

Submitted (21-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Feb 3, 2002 this sequence version replaced gi:18151062.

----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Web site: http://genome.wustl.edu/gsc  
 Contact: saplens@wustl.wustl.edu  
 Summary Statistics  
 Center project name: H\_NH0711J03  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>  
 VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-64A1, 2000 bp overlap; the clone sequenced to the right is RP11-397E7, 2000 bp overlap. Actual start of this clone is at base position 1995 of RP11-64A1; actual end is at base position 108789 of RP11-397E7.

#### FEATURES

##### source

1. 23618  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="4"  
 /map="4"

```

repeat_region      /clone="RP11-711J3"
repeat_region      /clone_id="RPCT-11"
repeat_region      595..720
repeat_region      /rpt_family="Alu"
repeat_region      789..1014
repeat_region      /rpt_family="Alu"
repeat_region      1039..1072
repeat_region      /rpt_family="(TAA)n"
repeat_region      1734..1936
repeat_region      /rpt_family="MIR"
repeat_region      2029..2175
repeat_region      /rpt_family="L1"
misc_feature        2339..2706
misc_feature        /note="similar to Sus scrofa EST BE031975 (NID:98326984)"
misc_feature        2484..2656
misc_feature        /note="match to EST BE181226 (NID:98660402)"
misc_feature        2499..2656
misc_feature        /note="similar to Mus musculus EST BB613812 (NID:916454310)"
repeat_region      2740..2892
repeat_region      /rpt_family="MIR"
repeat_region      3617..3920
repeat_region      /rpt_family="Alu"
repeat_region      4602..4890
repeat_region      /rpt_family="Alu"
repeat_region      4892..4914
repeat_region      /rpt_family="AT-rich"
repeat_region      5459..5624
repeat_region      /rpt_family="Alu"
misc_feature        6280..6368
misc_feature        /note="match to EST BE181226 (NID:98660402)"
misc_feature        6280..6368
misc_feature        /note="similar to Mus musculus EST BB613812 (NID:916454310)"
repeat_region      6811..7098
repeat_region      /rpt_family="Alu"
repeat_region      7515..7694
repeat_region      /rpt_family="MIR"
misc_feature        7794..7912
misc_feature        /note="similar to Mus musculus EST BB613812 (NID:916454310)"
misc_feature        7794..7885
misc_feature        /note="match to EST BE181226 (NID:98660402)"
repeat_region      8198..8513
repeat_region      /rpt_family="L1"
repeat_region      8514..8910
repeat_region      /rpt_family="MALR"
repeat_region      8911..9173
repeat_region      /rpt_family="L1"
repeat_region      9184..9592
repeat_region      /rpt_family="MALR"
repeat_region      9873..10183
repeat_region      /rpt_family="Alu"
repeat_region      10157..10266
repeat_region      /rpt_family="GA-rich"
repeat_region      10822..11165
repeat_region      /rpt_family="MALR"
repeat_region      11242..11261
repeat_region      /rpt_family="(TTTG)n"
repeat_region      11243..11530
repeat_region      /rpt_family="Alu"
repeat_region      11690..11833
repeat_region      /rpt_family="GA-rich"
repeat_region      11870..11929
repeat_region      /rpt_family="L1"
repeat_region      11916..11938
repeat_region      /rpt_family="AT-rich"
repeat_region      12632..12696
repeat_region      /rpt_family="L1"
repeat_region      12842..12935
repeat_region      /rpt_family="L2"
repeat_region      13129..13424
repeat_region      /rpt_family="Alu"

```

```

repeat_region      13401..13424
                    /rpt_family="(A)n"
repeat_region      13607..13641
                    /rpt_family="(A)n"
repeat_region      13642..13952
                    /rpt_family="(A)n"
repeat_region      13953..14359
                    /rpt_family="(A)n"
repeat_region      14392..14416
                    /rpt_family="(A)n"
repeat_region      14562..14588
                    /rpt_family="(A)n"
repeat_region      15001..15392
                    /rpt_family="(A)n"
repeat_region      16436..16604
                    /rpt_family="(A)n"
repeat_region      16895..17214
                    /rpt_family="(A)n"
repeat_region      17186..17231
                    /rpt_family="(A)n"
repeat_region      17275..17373
                    /rpt_family="(A)n"
repeat_region      17344..17657
                    /rpt_family="(A)n"
repeat_region      17713..17824
                    /rpt_family="(A)n"
repeat_region      18000..18165
                    /rpt_family="(A)n"
repeat_region      18235..18384
                    /rpt_family="(A)n"
repeat_region      18385..18693
                    /rpt_family="(A)n"
repeat_region      18666..18695
                    /rpt_family="(A)n"
repeat_region      18694..18807
                    /rpt_family="(A)n"
repeat_region      18881..19188
                    /rpt_family="(A)n"
repeat_region      19189..19562
                    /rpt_family="(A)n"
repeat_region      19563..19626
                    /rpt_family="(A)n"

```

Query Match 33.2% Score 377; DB 9; Length 23618;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-104;  
 Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 ATGAGAGGCAATGTTCTTCAGAGCTCAGCTCCCTCCCAAGATTCAAGAGAGAGCTG 60
23603 ATGAGAGGCAATGTTCTTCAGAGCTCAGCTCCCTCCCAAGATTCAAGAGAGAGCTG 23544

61 CCAGTGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
23543 CCAGTGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 23484

121 ATGATGGAGCTGCTCATGTTCTCTTGGAGATGTTCCGAGAGATCCGAGAGCTGAGT 180
23483 ATGATGGAGCTGCTCATGTTCTCTTGGAGATGTTCCGAGAGATCCGAGAGCTGAGT 23444

181 CACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
23423 CACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 23364

241 TTATAGGCTTATCTCTGAGCAGATTAGCTTTCTGAGAGCAGTCCAGAGCTATGCTGT 300
23363 TTATAGGCTTATCTCTGAGCAGATTAGCTTTCTGAGAGCAGTCCAGAGCTATGCTGT 23304

301 CTCATCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
23303 CTCATCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 23244

361 GGAGATATGAGATCTCAG 377
23243 GGAGATATGAGATCTCAG 23227

```

RESULT 2  
 AC093827/c  
 LOCUS  
 DEFINITION  
 AC093827 AC016973  
 VERSION  
 AC093827.3 GI:16328304  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 192263)  
 2 (bases 1 to 192263)  
 3 (bases 1 to 192263)  
 4 (bases 1 to 192263)  
 5 (bases 1 to 192263)  
 6 (bases 1 to 192263)  
 7 (bases 1 to 192263)  
 8 (bases 1 to 192263)  
 9 (bases 1 to 192263)  
 10 (bases 1 to 192263)  
 11 (bases 1 to 192263)  
 12 (bases 1 to 192263)  
 13 (bases 1 to 192263)  
 14 (bases 1 to 192263)  
 15 (bases 1 to 192263)  
 16 (bases 1 to 192263)  
 17 (bases 1 to 192263)  
 18 (bases 1 to 192263)  
 19 (bases 1 to 192263)  
 20 (bases 1 to 192263)  
 21 (bases 1 to 192263)  
 22 (bases 1 to 192263)  
 23 (bases 1 to 192263)  
 24 (bases 1 to 192263)  
 25 (bases 1 to 192263)  
 26 (bases 1 to 192263)  
 27 (bases 1 to 192263)  
 28 (bases 1 to 192263)  
 29 (bases 1 to 192263)  
 30 (bases 1 to 192263)  
 31 (bases 1 to 192263)  
 32 (bases 1 to 192263)  
 33 (bases 1 to 192263)  
 34 (bases 1 to 192263)  
 35 (bases 1 to 192263)  
 36 (bases 1 to 192263)  
 37 (bases 1 to 192263)  
 38 (bases 1 to 192263)  
 39 (bases 1 to 192263)  
 40 (bases 1 to 192263)  
 41 (bases 1 to 192263)  
 42 (bases 1 to 192263)  
 43 (bases 1 to 192263)  
 44 (bases 1 to 192263)  
 45 (bases 1 to 192263)  
 46 (bases 1 to 192263)  
 47 (bases 1 to 192263)  
 48 (bases 1 to 192263)  
 49 (bases 1 to 192263)  
 50 (bases 1 to 192263)  
 51 (bases 1 to 192263)  
 52 (bases 1 to 192263)  
 53 (bases 1 to 192263)  
 54 (bases 1 to 192263)  
 55 (bases 1 to 192263)  
 56 (bases 1 to 192263)  
 57 (bases 1 to 192263)  
 58 (bases 1 to 192263)  
 59 (bases 1 to 192263)  
 60 (bases 1 to 192263)  
 61 (bases 1 to 192263)  
 62 (bases 1 to 192263)  
 63 (bases 1 to 192263)  
 64 (bases 1 to 192263)  
 65 (bases 1 to 192263)  
 66 (bases 1 to 192263)  
 67 (bases 1 to 192263)  
 68 (bases 1 to 192263)  
 69 (bases 1 to 192263)  
 70 (bases 1 to 192263)  
 71 (bases 1 to 192263)  
 72 (bases 1 to 192263)  
 73 (bases 1 to 192263)  
 74 (bases 1 to 192263)  
 75 (bases 1 to 192263)  
 76 (bases 1 to 192263)  
 77 (bases 1 to 192263)  
 78 (bases 1 to 192263)  
 79 (bases 1 to 192263)  
 80 (bases 1 to 192263)  
 81 (bases 1 to 192263)  
 82 (bases 1 to 192263)  
 83 (bases 1 to 192263)  
 84 (bases 1 to 192263)  
 85 (bases 1 to 192263)  
 86 (bases 1 to 192263)  
 87 (bases 1 to 192263)  
 88 (bases 1 to 192263)  
 89 (bases 1 to 192263)  
 90 (bases 1 to 192263)  
 91 (bases 1 to 192263)  
 92 (bases 1 to 192263)  
 93 (bases 1 to 192263)  
 94 (bases 1 to 192263)  
 95 (bases 1 to 192263)  
 96 (bases 1 to 192263)  
 97 (bases 1 to 192263)  
 98 (bases 1 to 192263)  
 99 (bases 1 to 192263)  
 100 (bases 1 to 192263)

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
 The RPI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E.,

Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP11-711J3, the clone sequenced to the right is RP11-168E22. Actual start of this clone is at position 1 of RP11-397E7, actual end is at base position 192263 of RP11-397E7.

Data from AC079237 and AC093779 was used to finish this clone, AC093827. Polymorphisms have been identified between AC079237 and AC093827.

# FEATURES

The sequence of AC016973 has been incorporated into AC093827.

source  
1. 192263  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/map="4"  
/clone="RP11-397E7"  
/clone\_1lb="RPCT-11"  
1. 194  
/rpt\_family="MIR"  
misc\_feature  
1553. 1820  
/note="similar to EST BE151388 (NID:g8614109)"  
repeat\_region  
3422. 3792  
/rpt\_family="MALR"  
3917. 3942  
/rpt\_family="(T)n"  
4460. 4772  
/rpt\_family="Alu"  
4750. 4772  
/rpt\_family="(A)n"  
6150. 6469  
/rpt\_family="MALR"  
6470. 6781  
/rpt\_family="Alu"  
6754. 6795  
/rpt\_family="(GAAA)n"  
6823. 7127  
/rpt\_family="Alu"  
7245. 7324  
/rpt\_family="L1"  
7386. 7651  
/rpt\_family="Alu"  
7721. 7831  
/rpt\_family="L1"  
7853. 8154  
/rpt\_family="Alu"  
7967. 7987  
/rpt\_family="AT\_rich"  
8134. 8165  
/rpt\_family="AT\_rich"  
8269. 8410  
/rpt\_family="L1"  
9108. 9235  
/rpt\_family="MIR"  
9966. 10007  
/rpt\_family="MIR"  
10729. 11025  
/rpt\_family="Alu"  
11003. 11025  
/rpt\_family="(A)n"  
11207. 11818  
/rpt\_family="L1"  
11808. 11850  
/rpt\_family="AT\_rich"  
11851. 12151

misc\_feature  
12109. 12705  
/note="similar to EST BG619594 (NID:g13670965)"  
repeat\_region  
12132. 12151  
/rpt\_family="(A)n"  
12788. 12884  
/rpt\_family="Alu"  
13357. 13649  
/rpt\_family="Alu"  
13735. 14280  
/rpt\_family="L1"  
14376. 14499  
/rpt\_family="Alu"  
14500. 15875  
/rpt\_family="L1"  
15857. 15880  
/rpt\_family="AT\_rich"  
15876. 16047  
/rpt\_family="Alu"  
16086. 16231  
/rpt\_family="Alu"  
16254. 16690  
/rpt\_family="L2"  
21991. 22029  
/rpt\_family="AT\_rich"  
22069. 22413  
/rpt\_family="L2"  
22460. 22610  
/rpt\_family="Alu"  
23018. 23063  
/note="similar to EST B1522604 (NID:g15347396)"  
23034. 23552  
/note="similar to EST BG707405 (NID:g13983721)"  
23040. 23737  
/note="similar to EST BG612893 (NID:g13664264)"  
23042. 23907  
/note="similar to EST B1549585 (NID:g15436897)"  
23042. 23852  
/note="similar to EST BG779925 (NID:g14050242)"  
23042. 23791  
/note="similar to EST BG401954 (NID:g13295402)"  
23042. 23696  
/note="similar to EST BG20731 (NID:g13999918)"  
23042. 23491  
/note="similar to EST BG530513 (NID:g13522050)"  
23043. 23763  
/note="similar to EST BF700412 (NID:g11985820)"  
23043. 23273  
/note="similar to EST BG771989 (NID:g14082642)"  
23044. 23864  
/note="similar to EST BG528480 (NID:g13520017)"  
23053. 23874  
/note="similar to EST B1838666 (NID:g15950216)"  
23055. 23873  
/note="similar to EST BG622780 (NID:g14170367)"  
23056. 23907  
/note="similar to EST BG680207 (NID:g13911604)"  
23056. 23907

Query Match 33.2%; Score 377; DB 9; Length 192263;  
Best Local Similarity 100.0%; Pred. No. 6.9e-104; Indels 0; Gaps 0;  
Matches 377; Conservative 0; Mismatches 0;

QY 1 ATGAGAGCCAAATGTTTCAGAGAGCTGACCTGCGCCGCAACAGTTCCAGAGAGAGAGCTG 60  
DB 1985 ATGAGAGCCAAATGTTTCAGAGAGCTGACCTGCGCCGCAACAGTTCCAGAGAGAGAGCTG 1926  
QY 61 CCAGTGGAGCTGAGAGTCAAGAACTGAGAGCTGTTTTCACAGTGTGTCAGCTGTG 120  
DB 1995 CCAGTGGAGCTGAGAGTCAAGAACTGAGAGCTGTTTTCACAGTGTGTCAGCTGTG 1866  
QY 121 ATGATGGAGCTGCTCATGCTCTTTCAGAGAGTTCGTTGAGAGTCCGAGAGCTGTGTGTCG 180

Db 1865 ATGATGGGCGTGTCTATGTTCTTTGGAGTTCCTGGAGATCCGGAAGCTGTGCTG 1806  
Oy 181 CACATCAGAGAGACCTGGGGGAGTGTGCTGTGGAGTCTGTGCGAGTTGGGCTATGCT 240  
Db 1805 CACATCAGAGAGACCTGGGGGAGTGTGCTGTGGAGTCTGTGCGAGTTGGGCTATGCT 1746  
Oy 241 TTTACAGTTTATCTCTCTGCGCATTAGCTTTTCTGGAAGCCAGTCCAGTATTGCTGT 300  
Db 1745 TTTACAGTTTATCTCTCTGCGCATTAGCTTTTCTGGAAGCCAGTCCAGTATTGCTGT 1686  
Oy 301 CTCATCATGGGCTGTGCTGCGCGGGGACCATCTCTAACATTTTCACTTCTGGTTGAT 360  
Db 1685 CTCATCATGGGCTGTGCTGCGCGGGGACCATCTCTAACATTTTCACTTCTGGTTGAT 1626  
Oy 361 GGAGATATGATCTCAG 377  
Db 1625 GGAGATATGATCTCAG 1609  
LOCUS 1865  
DEFINITION Homo sapiens chromosome 18 clone RP11-819K4 map 18, LOW-PASS  
ACCESSION AC099847 65268 bp DNA linear HTG 22-NOV-2001  
VERSION AC099847.1 GI:17047210  
KEYWORDS HTG; HTGS PHASE0.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 65268)  
Birren, B., Linton, L., Nussbaum, C. and Lander, E.  
Homo sapiens chromosome 18, clone RP11-819K4  
Unpublished  
2 (bases 1 to 65268)  
Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barra, N., Bastien, V., Boguslavsky, L., Bouckhalter, B.,  
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
Chepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,  
Cooke, P., Dearliano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S.,  
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Gilde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,  
Lamatzes, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,  
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. S., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
TITLE Whitehead Institute/MIT Center for Genome Research  
JOURNAL Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Web site: http://www-seq.wi.mit.edu  
Center code: WIRB  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L13211  
Center clone name: 819\_K\_4

\* NOTE: This record contains 81 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1 727: contig of 727 bp in length  
\* 728 827: gap of 100 bp  
\* 828 1547: contig of 720 bp in length  
\* 1548 1647: gap of 100 bp  
\* 1648 2369: contig of 722 bp in length  
\* 2370 2469: gap of 100 bp  
\* 2470 3148: contig of 679 bp in length  
\* 3149 3248: gap of 100 bp  
\* 3249 3922: contig of 674 bp in length  
\* 3923 4022: gap of 100 bp  
\* 4023 4708: contig of 686 bp in length  
\* 4709 4808: gap of 100 bp  
\* 4809 5512: contig of 704 bp in length  
\* 5513 5612: gap of 100 bp  
\* 5613 6320: contig of 708 bp in length  
\* 6321 6420: gap of 100 bp  
\* 6421 7119: contig of 699 bp in length  
\* 7120 7219: gap of 100 bp  
\* 7220 7943: contig of 724 bp in length  
\* 7944 8043: gap of 100 bp  
\* 8044 8780: contig of 737 bp in length  
\* 8781 8880: gap of 100 bp  
\* 8881 9602: contig of 722 bp in length  
\* 9603 9702: gap of 100 bp  
\* 9703 10429: contig of 727 bp in length  
\* 10430 10529: gap of 100 bp  
\* 10530 11235: contig of 706 bp in length  
\* 11236 11335: gap of 100 bp  
\* 11336 12030: contig of 695 bp in length  
\* 12031 12130: gap of 100 bp  
\* 12131 12822: contig of 692 bp in length  
\* 12823 12922: gap of 100 bp  
\* 12923 13644: contig of 722 bp in length  
\* 13645 13744: gap of 100 bp  
\* 13745 14466: contig of 722 bp in length  
\* 14467 14566: gap of 100 bp  
\* 14567 15297: contig of 731 bp in length  
\* 15298 15397: gap of 100 bp  
\* 15398 16109: contig of 712 bp in length  
\* 16110 16209: gap of 100 bp  
\* 16210 16935: contig of 726 bp in length  
\* 16936 17035: gap of 100 bp  
\* 17036 17749: contig of 714 bp in length  
\* 17750 17849: gap of 100 bp  
\* 17850 18542: contig of 693 bp in length  
\* 18543 18642: gap of 100 bp  
\* 18643 19341: contig of 699 bp in length  
\* 19342 19441: gap of 100 bp  
\* 19442 20144: contig of 703 bp in length  
\* 20145 20244: gap of 100 bp  
\* 20245 20963: contig of 719 bp in length  
\* 20964 21063: gap of 100 bp  
\* 21064 21779: contig of 716 bp in length  
\* 21780 21879: gap of 100 bp  
\* 21880 22596: contig of 717 bp in length  
\* 22597 22696: gap of 100 bp  
\* 22697 23361: contig of 665 bp in length  
\* 23362 23461: gap of 100 bp  
\* 23462 24186: contig of 725 bp in length  
\* 24187 24286: gap of 100 bp  
\* 24287 25016: contig of 730 bp in length  
\* 25017 25116: gap of 100 bp

25117 25835: contig of 719 bp in length  
25836 25935: gap of 100 bp  
25936 26607: contig of 672 bp in length  
26608 26707: gap of 100 bp  
26708 27297: contig of 590 bp in length  
27298 27397: gap of 100 bp  
27398 28099: contig of 702 bp in length  
28100 28199: gap of 100 bp  
28200 28896: contig of 697 bp in length  
28897 28996: gap of 100 bp  
28997 29707: contig of 711 bp in length  
29708 29807: gap of 100 bp  
29808 30519: contig of 712 bp in length  
30520 30619: gap of 100 bp  
30620 31348: contig of 729 bp in length  
31349 31448: gap of 100 bp  
31449 32177: contig of 729 bp in length  
32178 32277: gap of 100 bp  
32278 33007: contig of 730 bp in length  
33008 33107: gap of 100 bp  
33108 33808: contig of 701 bp in length  
33809 33908: gap of 100 bp  
33909 34618: contig of 710 bp in length  
34619 34718: gap of 100 bp  
34719 35387: contig of 669 bp in length  
35388 35487: gap of 100 bp  
35488 36158: contig of 671 bp in length  
36159 36258: gap of 100 bp  
36259 36957: contig of 699 bp in length  
36958 37057: gap of 100 bp  
37058 37764: contig of 707 bp in length  
37765 37864: gap of 100 bp  
37865 38561: contig of 697 bp in length  
38562 38661: gap of 100 bp  
38662 39375: contig of 714 bp in length  
39376 39475: gap of 100 bp  
39476 40191: contig of 716 bp in length  
40192 40291: gap of 100 bp  
40292 41017: contig of 726 bp in length  
41018 41117: gap of 100 bp  
41118 41764: contig of 647 bp in length  
41765 41864: gap of 100 bp  
41865 42593: contig of 729 bp in length  
42594 42693: gap of 100 bp  
42694 43356: contig of 663 bp in length  
43357 43456: gap of 100 bp  
43457 44158: contig of 702 bp in length  
44159 44258: gap of 100 bp  
44259 44951: contig of 693 bp in length  
44952 45051: gap of 100 bp  
45052 45766: contig of 715 bp in length  
45767 45866: gap of 100 bp  
45867 46581: contig of 715 bp in length  
46582 46681: gap of 100 bp  
46682 47400: contig of 719 bp in length  
47401 47500: gap of 100 bp  
47501 48234: contig of 734 bp in length  
48235 48334: gap of 100 bp  
48335 49057: contig of 723 bp in length  
49058 49157: gap of 100 bp  
49158 49877: contig of 720 bp in length  
49878 49977: gap of 100 bp  
49978 50699: contig of 722 bp in length  
50700 50799: gap of 100 bp  
50800 51494: contig of 695 bp in length  
51495 51594: gap of 100 bp  
51595 52298: contig of 704 bp in length  
52299 52398: gap of 100 bp  
52399 53096: contig of 698 bp in length  
53097 53196: gap of 100 bp  
53197 53903: contig of 707 bp in length  
53904 54003: gap of 100 bp  
54004 54708: contig of 705 bp in length

\* 54709 54808: gap of 100 bp  
\* 54809 55524: contig of 716 bp in length

Query Match  
Best Local Similarity 33.1%; Score 375.4; DB 2; Length 65268;  
Matches 376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAGACCAATGTTTCACAGCTCAGCTCAGCTCCTCCACAGTTTCAGAGAGAGCTG 60  
50030 ATGAGAGCCATGTTTCACAGCTCAGCTCAGCTCCTCCACAGTTTCAGAGAGAGCTG 50089

61 CCAGTGGAGCTGAGAGTGAATGAACTGAGAGCTGTTTACAGTGTGCTGCTG 120  
50090 CCAGTGGAGCTGAGAGTGAATGAACTGAGAGCTGTTTACAGTGTGCTGCTG 50149

121 ATGATGGAGCTGCTCATGTTCTCTTGGAGATGTTCCGTGAGATCCGAGAGCTGCTG 180  
50150 ATGATGGAGCTGCTCATGTTCTCTTGGAGATGTTCCGTGAGATCCGAGAGCTGCTG 50209

181 CACATCAGAGACCTCGGGGCAATTCCTGTGGAGCTGCTGCTGCTGCTGCT 240  
50210 CACATCAGAGACCTCGGGGCAATTCCTGTGGAGCTGCTGCTGCTGCTGCTGCT 50269

241 TTTACAGCTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
50270 TTTACAGCTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 50329

301 CTCATCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
50330 CTCATCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 50389

361 GGAGATATGATCTCAG 377

Db 50390 GGAGATATGATCTCAG 50406

RESULT 4  
AR033870  
LOCUS AR033870 2263 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5869265.  
ACCESSION AR033870  
VERSION AR033870.1 GI:5949475  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2263)  
AUTHORS Dawson,P.A.  
TITLE Illeal bile acid transporter compositions and methods  
JOURNAL Patent: US 5869265-A 1 09-FEB-1999;  
FEATURES  
source location/Qualifiers  
1..2263  
/organism="unknown"

BASE COUNT 672 a 451 c 476 g 664 t

ORIGIN

Query Match 28.3%; Score 320.4; DB 6; Length 2263;  
Best Local Similarity 60.8%; Pred. No. 9e-87;  
Matches 522; Conservative 0; Mismatches 336; Indels 0; Gaps 0;

80 ATGGAACCTGAGCTGCTTTCACAGTGTGCTCAGTGTGATGATGAGGCTGCTCATGT 139  
188 AGGCCATCTTACGCTGTGATGAGACCTGCTCACAATCTCTTACCTTGTGATGT 247

140 TCTCTTGGAGATGCTCGTGAATCCGAAAGCTGTGCTCACAATCAGAGACCTGCG 199  
248 TTTTCATGAGGATGCAATGTGAATCCCAAGTTTTCGAGACCTAAGGGGCGCATGG 307

200 GCATTCGTGAGAGCTGCTGCTGCAATTTGAGCTGATGCTTTTACAGCTTATCTCTGG 259  
308 GCATGCTGTGAGGCTTCTGCTGTGATTTGGAATCATGCTCTCAGAGTTTTCGCTGT 367

260 CCATTAGCTTTTCTGTAAGCAATCAAGCTATTGCTGCTATCATCATGAGGCTGCTGCC 319



Db 368 CGGTGGCCCTTGGCATCTCCCGAGGCAAGCTGTGTGTGTATTCAGAGGTTGCTCC 427  
Qy 320 CGGGGGGACCATCTCTAATATTTTACCTTGTGGTGTATGAGATATGATCTCACA 379  
Db 428 CTGGAGGAATGCTCTCAATATCTTAGGCTTATTTGGTGTATGGCAATGAGCTCAGC 487  
Qy 380 TCAGATATGACAACTGTTCACACGCGGCGCTGGAGATGATGCACTGTGCAATTATC 439  
Db 488 TTAGATATACCACTGTCTCAGCGCTGTGCTTGGATGATGATGCTTGTGCTTCA 547  
Qy 440 TCTACACCTGTGTCTGAGATCTTTCAGAGATCTGACATCTCTTATCAGAACATAGAA 499  
Db 548 TCTATACCAAGATGAGGTGATCTCAGGGAAGATGTATCTTATATGACAGATTTGCA 607  
Qy 500 TTACCTTGTGTGTGCTGACATCTCTGTGGCTTGTGTGTATGATGATTAATGATGCG 559  
Db 608 CTCTCTGTGTGTCTTGTATCTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 667  
Qy 560 CAAACATCCAAATATCTCTCAAGATTTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 619  
Db 668 CCGAAAGGCAAGATCTATCTTAATTTGATTCATGCGAGTGTCAATCTCATTTGTC 727  
Qy 620 TGTGTGAGT 679  
Db 728 TCATGCTGT 787  
Qy 680 TGACCATCATTTTCTCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 739  
Db 788 GATATATAGGAACATATATCTTATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 847  
Qy 740 TTACCAACAGCTTGTGGGAAAGTGTGAGAGCAATTTCTTGAATCTGAGCTCAGATA 799  
Db 848 TTGCTGT 907  
Qy 800 TTCAGATGTGATCAACATGTCTCAGTATCTTCTCTGCTGTGAGCACTTGTGTGAGAT 859  
Db 908 CTCAGCTGT 967  
Qy 860 TGAATTTTCCACTGTGCTATGAGATCTTTCAGCTGATATGATGATTTCTTATTTGTGAG 919  
Db 968 TCACCTTCCCTCTATCTACAGCATCTTCCAGATGCTGTGTGTGTGTGTGTGTGTGTGT 1027  
Qy 920 CATATCAGAGCTACAGA 937  
Db 1028 CTTATGTGCATACAGA 1045  
RESULT 5  
LOCUS 132744 2263 bp DNA linear PAT 06-FEB-1997  
DEFINITION Sequence 1 from patent US 5589358.  
ACCESSION 132744 GI:1823535  
VERSION 132744.1  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2263)  
AUTHORS Dawson, P.A.  
TITLE Ileal bile acid transporter compositions and methods  
JOURNAL Patent: US 5589358-A 1 31-DEC-1996;  
FEATURES  
source location/Qualifiers  
1..2263  
/organism="unknown"  
BASE COUNT 672 a 451 c 476 g 664 t  
ORIGIN  
Query Match 28.3%; Score 320.4; DB 6; length 2263;  
Best Local Similarity 60.8%; Pred. No. 96-87;  
Matches 522; Conservative 0; Mismatches 336; Indels 0; Gaps 0;  
Qy 80 ATGAAACCTGAGCTGTTTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 139

Db 188 AGCGCATCTCAGGCGGTGTATGAGCAACGCTGTCAATCTCTGATGATGT 247  
Qy 140 TCTCTTGGATGTTCCTGTGAGATCCGGAAGCTGTGTGTGTGTGTGTGTGTGTGTGT 199  
Db 248 TTTCCATGGGGGTGCAATGT 307  
Qy 200 GCATTTGCTGTGAGCTGT 259  
Db 308 GCATGT 367  
Qy 260 CCATTAAGCTTTCTGTGAAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 319  
Db 368 CGGTGGCCCTTGGCATCTCCCGAGGCAAGCTGTGTGTGTATTCAGAGGTTGCTCC 427  
Qy 320 CGGGGGGACCATCTCTAATATTTTACCTTGTGGTGTATGAGATATGATCTCACA 379  
Db 428 CTGGAGGAATGCTCTCAATATCTTAGGCTTATTTGGTGTATGGCAATGAGCTCAGC 487  
Qy 380 TCAGATATGACAACTGTTCACACGCGGCGCTGGAGATGATGCACTGTGCAATTATC 439  
Db 488 TTAGATATACCACTGTCTCAGCGCTGTGCTTGGATGATGATGCTTGTGCTTCA 547  
Qy 440 TCTACACCTGTGTCTGAGATCTTTCAGAGATCTGACATCTCTTATCAGAACATAGAA 499  
Db 548 TCTATACCAAGATGAGGTGATCTCAGGGAAGATGTATCTTATATGACAGATTTGCA 607  
Qy 500 TTACCTTGTGTGTGCTGACATCTCTGTGGCTTGTGTGTATGATGATTAATGATGCG 559  
Db 608 CTCTCTGTGTGTCTTGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 667  
Qy 560 CAAACATCCAAATATCTCTCAAGATTTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 619  
Db 668 CCGAAAGGCAAGATCTATCTTAATTTGATTCATGCGAGTGTCAATCTCATTTGTC 727  
Qy 620 TGTGTGAGT 679  
Db 728 TCATGCTGT 787  
Qy 680 TGACCATCATTTTCTCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 739  
Db 788 GATATATAGGAACATATATCTTATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 847  
Qy 740 TTACCAACAGCTTGTGGGAAAGTGTGAGAGCAATTTCTTGAATCTGAGCTCAGATA 799  
Db 848 TTGCTGT 907  
Qy 800 TTCAGATGTGATCAACATGTCTCAGTATCTTCTCTGCTGTGAGCACTTGTGTGAGAT 859  
Db 908 CTCAGCTGT 967  
Qy 860 TGAATTTTCCACTGTGCTATGAGATCTTTCAGCTGATATGATGATTTCTTATTTGTGAG 919  
Db 968 TCACCTTCCCTCTATCTACAGCATCTTCCAGATGCTGTGTGTGTGTGTGTGTGTGTGT 1027  
Qy 920 CATATCAGAGCTACAGA 937  
Db 1028 CTTATGTGCATACAGA 1045  
RESULT 6  
LOCUS CGU02028 2263 bp mRNA linear ROD 06-JUN-1994  
DEFINITION Cricetus griseus Na+ dependent ileal bile acid transporter mRNA,  
complete cds.  
ACCESSION U02028  
VERSION U02028.1 GI:455032  
KEYWORDS  
SOURCE Chinese hamster.  
ORGANISM Cricetus griseus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Cricetus.

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
1 (bases 1 to 2263)	Wong, M.H., Oelkers, P., Craddock, A.L. and Dawson, P.A.	Expression cloning and characterization of the hamster ileal sodium-dependent bile acid transporter	J. Biol. Chem.	269 (2), 1340-1347 (1994)	94117449	2 (bases 1 to 2263)	Dawson, P.A.	Direct Submission	Submitted (22-SEP-1993)	Paul A. Dawson, Dept Medicine/Section Gastroenterology, Bowman Gray School of Medicine, Wake Forest University, Medical Center Boulevard, Winston-Salem, NC, 27157, USA	Location/Qualifiers
1..2263		/organism="Cricetulus griseus"									
/db_xref="taxon:10029"		/clone="clone PIBAT(44-1)"									
/tissue_type="ileum"		/clone_lib="hamster ileal cDNA expression library"									
/note="author cites additional common name: golden Syrian hamster"											
109..1155		/codon_start=1									
/product="Na+ dependent ileal bile acid transporter"		/protein_id="AA18640.1"									
/db_xref="GI:455033"											
/translation="MDNSSICNPNTICEGSDSCIAESPENFALLSVNLTLLALLVMSFGNVELHKEGLRLRPWGIYVGLCPGFMPLGFLVSAFGILPYQAVVLLIQSGCGTASNLIAWVDGMDLVSMTSGTLLALGMPCLFLYTMWVDSGLIYVYDSCTGVALVIVPSIGVMVNMHPKATILIKGISIACAILIYAVVGGILYOSAMTIERKMLITGITIPIYNGYGLGFLARLDAQPMYRCKTVLALEGLONTQLCSTYVLSFSEPDNLVTFPLIYISFQIAPALILGAVYVVKCHGNNTLEKTDNEPSPSSFOETNRKGFQPEDEK"											
109..1152		/product="Na+ dependent ileal bile acid transporter"									
BASE COUNT	672 a	451 c	476 g	664 t							
ORIGIN											
Query Match	28.3%	Score 320.4	DB 10	Length 2263							
Best Local Similarity	60.8%	Pred. No. 9e-87									
Matches 522	Conservative	0	Mismatches 336	Indels 0	Gaps 0						
QY	80	ATGGAACCTGGAGCTCGTTTTCACAGTGTGTCCACTGTGATGATGGGCGTCTCACT	139								
DB	188	ACGCCATCTCAGCGCTGTGTGATGAGACACCGTCTCCACATCTCTACCTCTTGGAGT	247								
QY	140	TCTCTTGGAGTGTCCGTGAGATCCGGAAGCTGTGTGCGACATCGAGAACCTCGGG	199								
DB	248	TTTCCATGGGGTGCATGTGGAACCTCCACAGTTTCTGGAGACCTTAAGCGCGCCATGG	307								
QY	200	GCATTGCTGTGGAGCTGTGCGACAGTTTGGCTCATGCTCTTTACAGCTTATCTCTGG	259								
DB	308	GCATCGTGTGGGCTTCCCTGTGTGAGTTTGGATCATGCTCTCCACAGTTTGGTCTGT	367								
QY	260	CCATTAGCTTTCCTTGAAGCCAGTTCACAGTATTGCTGTCTATCATGGGCTGTGCC	319								
DB	368	CCGTGGCCTTTGGCATCTCCACAGTGCAGAGCTGTGTGTGTATTCAGGGTTGTGCC	427								
QY	320	CGGGGGGACACATCTCTAACATTTTCACTTCTGGTGTGATGAGGATATGAGTCTACGA	379								
DB	428	CTGGAGGAATGCTCTCAATATCTTAGCTTATTTGGTATGATGGGACATGAGACTTCAGC	487								
QY	380	TCAGTATGACAACTGTTCACACCGTGGCGCGCTGGGATGATGCGCACTCTGACATTTATC	439								
DB	488	TTAGATATACACCTGTCTCAATATCTTAGCTTATTTGGTATGATGGGACATGAGACTTCAGC	547								
QY	440	TCATACCTGTCTGTGAGTCTTACGACAGATCTCACATTCCTTATCAGAACATAGAA	499								
DB	548	TCATATACCAAGATGTGGGTGACTCAGGAGACGATGTGATTCCTTATGACAGCATTTGCA	607								
QY	500	TTACCTGTGTGCTGTGACATTTCTGTGTGCTTTGGTGTGTCTATGTGAATTACAGATGC	559								

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	source
Db																	
608	CTTCTGCTGTTGCTCTTGTATTTCCTGTTTCCATTTGGAAATGTAATGTAATCACAAATGGC																
Qy	560	CAAACAATCAAAATATTTCTCAAGATTGGGGCCGTTGTGGTGGGGTCTCTCTTG	619														
Db	668	CCCAAAAAGCAAAAGATATCTATTAATTTGGATCCATCGCAGGCGCAATTTCTCATTTCTC	727														
Qy	620	TGGTCGAGATTGCTCGAGTGTGTCCTCGCGCAAAAGATTTTGGAAATTCAGATCAACCTTC	679														
Db	728	TCATCGCTGTGTTGGAGGAATCTGTACCAAAAGTGCCTTGACCATTTGAACCCAGCTGT	787														
Qy	680	TGACCATCAGTTTCATCTTTCTTTGATTGGCCATGTCAAGGTTTTCCTGCGCACTTT	739														
Db	788	GGATTATAGGAACATATATCTATATAGCTGGCTACAGGCTGTGGGGTTTTCCTGCGTAGAA	847														
Qy	740	TTATCCCAACAGTCTTGGCAAAAGTGCAGAGCAATTTCTTAGAAATCTGAGCTCAGATA	799														
Db	848	TTGCTGTGTCAACCTGTGTACAGGCGCGAACAGTGTGCTTGGAAACCGGGTTTCGAGAAC	907														
Qy	800	TTGATGATGATCAGATCAGATGCTCAGTTATCTTTCACTGCTGAGACATTTGGTCCAGATGT	859														
Db	908	CTACGCTGTGTTCACCATTTGTGACGCTTTCTTCAGCCTTGAGAGCTCAACCTGTGT	967														
Qy	860	TGACTTTCCACCTGCGCCTTATGACCTCTTCAGCTGATAGATGATTTCTTATTTGTTGACG	919														
Db	968	TCACCTTCCCGCTCATCTACAGCATCTTCCAGATCGCCTTTGCGACAACTATTAGAG	1027														
Qy	920	CATATCAGAGCTACAGA	937														
Db	1028	CTTATGTGCAATACAGA	1045														
RESULT 7																	
Db7059																	
LOCUS	Db7059	974 bp	mRNA	linear	ROD 07-FEB-1999												
DEFINITION	House mouse; Musculus domesticus	acid transporter, partial cds.															
ACCESSION	Db7059.1	GI:1504059															
VERSION	1																
KEYWORDS	l1eal Na+-dependent bile acid transporter.																
SOURCE	Mus musculus	(strain:ICR) 8 weeks CDNA to mRNA.															
ORGANISM	Mus musculus																
REFERENCE	1	(bases 1 to 974)															
AUTHORS	Saeki,T.																
TITLE	Mouse l1eal Na+-dependent bile acid transporter cDNA: Partial CDS																
JOURNAL	Unpublished																
REFERENCE	2	(bases 1 to 974)															
AUTHORS	Saeki,T.																
TITLE	Direct Submission																
JOURNAL	Submitted (09-AUG-1996)																
REFERENCE	Department of Biological Resource Chemistry; Nakaragi, Shmugamo,																
AUTHORS	Saeki,T.																
TITLE	Submitted (09-AUG-1996)																
JOURNAL	Submitted (09-AUG-1996)																
REFERENCE	Submitted (09-AUG-1996)																
AUTHORS	Submitted (09-AUG-1996)																
TITLE	Submitted (09-AUG-1996)																

Query Match	Similarity	27.1%	Score	307.8	DB	10	Length	974
Matches	522	Conservative	0	Mismatches	357	Indels	0	Gaps
Db	80	ATGAAACCTGGAGCTCGTTTTCACAGTGTGTCCACTGTATGATGAGGCGTCAATG	139					
Qy	140	TCCTCTTGGAGATGTTCCGTGGAGATCCGGAAGCTGTGGTCGACATCAGAGAACCTTGGG	199					
Db	116	TTTCTATGAGGAGTCATGATGGAAAGTCCACAAAGTCTTCAGAGATATTAAGAACCATGAGG	175					
Qy	200	GCATGTCTGTGGGACCTGCTCTGCCAGTTTGGGCGTCATGCGCTTTTACAGCTTATCTCTGG	259					
Db	176	GTAATCTTGCTGGGCTTCTCTCTGTACGTTTGGATATATGCTCTTCACAGGCTTATCTCTGT	235					
Qy	260	CCATTAGCTTTTCTCTGGAAGCCAGTCCAAAGCTATTGCTGTTCATCATGAGGCTGTGCG	319					
Db	236	CTGTGGCCTGTGGCATCTCTCTGTACAGGCTGTAGTGGTGTAAATTAATGAGGTCTGTGCG	295					
Qy	320	CGGGGGGACCATCTCTTAACATTTTACCTTCTGGGTGATGAGAAATATGATCTTCAGCA	379					
Db	296	CTGGAGGAATCTGGCTCCAAATATCCCTGAGCTTATGGAATGATGAGGAGCAATGAGCTTCAGTG	355					
Qy	380	TCAGATATGACAACTGTTCACACGTGGCCGCGCTGGAGAAATGATGCACTGTCAATTTATC	439					
Db	356	TTAGATATACCACTGTCTCTCAACACTGTCTGGCTTGGATATGCTCTTGTGCTCTTGG	415					
Qy	440	TCATACCTGTGCTCTGTGAGTCTTCAGCAAGATCTGACCATTCCTTATCAGAACATAGAA	499					
Db	416	TCATACCAAGATATGAGGTGACTGGGAAACATGTTGATTCCTATGATATACATTTGTA	475					
Qy	500	TTACCTTGTGTGCTGTACCATTTCTGTGGCGCTTTGTGTCTATGTGAATTTACAGATGGC	559					
Db	476	TTTCTGTGTGCTCTGTATATCTGTGTTCTTTTGGAAATGTTTGAATTCACAAATGAGC	535					
Qy	560	CAAAACAATCCAAATTCATTCACAGATTTGGGCGCGTGTGTGGTGGGCGTCTCTTCTGG	619					
Db	536	CACAAAAACGAAAGTTATATCTTAAATTTGATTCATCAAGGTGTAATTCATCTTGTGC	595					
Qy	620	TGTTGCGAGTTGTGCTGTGTGCTCTGGCGGAAGATCTTGAATTCAGACATCACTCCCTTC	679					
Db	596	TCATGTAGCTGTGATGGAGGAATACGTACCAAGTGCCTGATTCATGAAACCAACTGT	655					
Qy	680	TGACCATCAGTTTCATCTTCTTGTGATTTGGGCATGTCAAGGCTTTTCTGTGGCACTTT	739					
Db	656	GGATTATAGGAACAAATATTCCTTATAGCTGTGCTACAGCTGTGGTCTTCTCTGTGGCTAGC	715					
Qy	740	TTAACCCACAGCTCTTGGCAAGAGTGCAGAGCAATTTCTTAAACTGGAGCTCAGAAATA	799					
Db	716	TAGCTGTGTACACCTGTGTACAGGTGCCAAACAGTAGCTTTGAAACTGGAATGCGAACA	775					
Qy	800	TTCAAGATGTCAATCCATGTCTCCAGTTATCTTTTCACTGTGACGACTTGTGTCAGATGT	859					
Db	776	CTCACTGTGTCTCCACCAATTTATACAGCTCTCTCTCCGCCAGAGATCTCAACTGTGTGT	835					
Qy	860	TGAGTTTCCCATTTGGCGCTATGAGACTCTTCCAGCTGTATAGATGATTTCTTATTTGTCAG	919					
Db	836	TCACCTTCCCACTCATCTATCTGTTTTCACCTGTCTTTTSCAGACGATTAATTAATGGA	895					
Qy	920	CATATCAGACGTATCAAGAGAGATTTGAAGAACAAACATG	958					
Db	896	TTTATGTCAATACAGGAATGTATGAGAAAAAATGATG	934					

SOURCE	Mus musculus (strain:ICR) 8 weeks male cDNA to mRNA.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1629)
AUTHORS	Saeeki,T., Matoba,K., Furukawa,H., Kirifuji,K., Kanamoto,R. and Iwami,K.
TITLE	Characterization, cDNA cloning, and functional expression of mouse ileal sodium-dependent bile acid transporter
JOURNAL	J. Biochem. 125 (4), 846-851 (1999)
MEDLINE	99203592
REFERENCE	2 (bases 1 to 1629)
AUTHORS	Saeeki,T.
TITLE	Direct Submission
JOURNAL	Submitted (07-APR-1997) Tohru Saeeki, Kyoto Prefectural University, Department of Biological Resource Chemistry; Nakaragi, Shinozono, Sakyo-ku, Kyoto 606, Japan [E-mail:t.saeeki@ids.kpu.ac.jp, Tel.:81-75-703-5663, Fax:81-75-703-5661]
FEATURES	location/qualifiers
source	1..1629
	/organism="Mus musculus"
	/strain="ICR"
	/db_xref="taxon:10090"
	/sex="male"
	/dev_stage="8 weeks"
	50..1096
	/codon_start=1
	/product="ISBr"
	/protein_id="BAAL19606.1"
	/db_xref="gi:1944179"
	/translation="MDNSVCPNPATVCEDGSQVPSPENPAIINTVMSTVLITLLAM VMEWGCVNVEHKFLGHIKKPMGIFVGLFCFPMPLPTGRILSLASGLIPQAVYLI MGCCPGGGSSIIIAWIDGDMDLSVNTKSTILLAMPCLPVLVYMWDSDTIYVA PYDGIISLVALVIPVSEFMGVNMHKKPOKAIILIKISIGVILIIVAVIGILYOS AMILEPKMLIGTIFPIAGYSIGEFLIALGOPLYRCRTVALFTGMONTOLCSTIVDS SESIEDNLVPTPLIYTVFQLVFAAVIILGIYVTRCYGNDAEPLEKTDNENDSRRE SFDRINKGFODEK"
CDS	
	BASE COUNT 451 a 315 c 341 g 522 t
ORIGIN	
Query Match	27.1%; Score 307.8; DB 10; Length 1629;
Best Local Similarity	59.4%; Pred. No. 6.4e-83;
Matches 522; Conservative	0; Mismatches 357; Indels 0; Gaps 0;
Dn	80 ATGAAACCTTGAGACTCGTTTTCACAGTGTGCACATGTGATGATGAGGGGCTCATGT 139
Dc	129 ATGCATTCTCAATACAGTAGTAGACATGTGCTCACCATCTCTTAGCCATGTGATGT 188
Dy	140 TCTCTTGGAGATGCCGTGGAATCCGGAAGCTGTGGTCCGACATCAGAGAACCTTGGG 199
Dd	189 TTTCATATGGGTGAGAATGTGGAAATGCCAAGTCTTCAGACAATAAAGAGACCATTGGG 248
Dq	200 GCATTGCTGTGGGACATGCTCTGCACAGTTTGGGCTCACAGSCCTTTACAGCTTAATCTCCGG 259
Df	249 GTATCTTGATGGCTTCTCTGTGTACAGTTTGAATCAATAGCCTCTCACAGGCTTTATCTGT 308
Dg	260 CCATTAGCTTTTCTCTGTAAGCCAGTCCAGCTAATGTGCTTCTCATCATCTGGGCTGTGCC 319
Dh	309 CTGTGGCCCTCTGGCATCCTCTCTGTACAGAGCTGTAGTGTCTAATTAATGGGTTGTGCC 368
Di	320 CGGGGGGACCAATCTTACAAATTTTACCTTCTGGGTGATGAGAAATATGGAATCTCAGA 379
Dj	369 CTGGAGGAATCTGGCTCCAATATCTCTGGGCTTATGGATAGATGGGAGACATGACCTCACTG 428
Dk	380 TCAGATATACAACTGTTCACACCGTGGCGGCTCGGGAATGATGCACTGTGCATTATTC 439
Dl	429 TTAGCATATACCACTGTTCACACCTGTGGCTTGGATATGATGCTTTTGGCTCTTTCG 488
Dm	440 TCTACACCTGTGCTCTGAGTCTTTCAGCAGAAATTCACACCTCTCTTATCAGAAATAGGA 499
Dn	489 TCTACACCAAAATGTGGGTTGATCTCGGAAAGATTTGATTCCTCATATGATAGATTGGTA 548
Do	500 TTACCTTGTTGTGCTGACCAATCCTCTGTGGCTTGTGTCTATATGCAATTACAGATGCC 559

Db 549 TTTCTGCTGCTCTCTGTTATTCCTGTTCCCTTGGAATGTTTGAATCACAATGTC 608  
Qy 560 CAAACCAATCCAAATCATTTCTCAAGATTGGGGCGGTTTGGGGGCTCCCTCTGG 619  
Db 609 CACAAAAGCGAATTATTAATTAATGATTCATCAAGGTGAATTTCAATTCATGTC 668  
Qy 620 TGGTGGCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679  
Db 669 TCATGCTGCTGATGGAAGAAATCTGTACCAAGTCCGTGATCATGAAACCAACTGT 728  
Qy 680 TGAACCATCACTTCACTTTCTTTGATGCGCATGTCACGGGTTTCTGCTGGCACTT 739  
Db 729 GGAATTATAGGAACAATATTCCTATAGCTGCTACAGCTGGGTTTCTTCTGCTGAGC 788  
Qy 740 TTACCCACAGCTTGGCAAGGTGCAAGCAATTTCTTAGAACTGAGCTCAGAAAT 799  
Db 789 TAGCTGCTCACTCTGTGACAGGTGCGCAACAGTACCTTGGAACTGGAATGCAAGAA 848  
Qy 800 TTCAGATGTCATCACCATGCTCCAGTTATCTTTCACTGCTGAGCACTTGTCCAGATG 859  
Db 849 CTCAGCTGCTGCTCCACATTTGATGACCTCTCTCTCCCGAGATCTCAACTGTGT 908  
Qy 860 TGAATTTCCCATGCTGCTTGAAGCTTTCCAGCTGATAGATGATTTCTTATTTGCGAG 919  
Db 909 TCACCTTCCCATCACTTCTACTGTTTCCAGCTCTCTTGTGACGATGATATATAGGA 968  
Qy 920 CATATCAGAGTACAGAGAGATGTAAGAACAAATG 958  
Db 969 TTTATGTCATACAGGAATGTTATGGAATAATGATG 1007

RESULT 9  
LOCUS OCSDBATRP 1116 bp RNA linear MM 12-OCT-1995  
DEFINITION O. cuniculus mRNA for ileal sodium-dependent bile acid transporter.  
ACCESSION 254357  
VERSION 254357.1 GI:1019395  
KEYWORDS ileal sodium-dependent bile acid transporter.  
SOURCE Oryctolagus cuniculus.  
ORGANISM Oryctolagus cuniculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
REFERENCE 1 (bases 1 to 1116)  
AUTHORS Stengelin, S., Apel, S., Becker, W., Maier, W., Rosenberger, J., Wess, G.  
and Kramer, W.  
TITLE Cloning of the rabbit ileal sodium-dependent bile acid transporter  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1116)  
AUTHORS Stengelin, S.  
TITLE Direct Submision  
JOURNAL Submitted (11-OCT-1995) Stengelin S., Hoechst Marion Roussel, TD  
Metabolism, Building H825, D-65926 Frankfurt am Main, Germany  
FEATURES  
source location/Qualifiers  
1..1116  
/organism="Oryctolagus cuniculus"  
/strain="New Zealand White"  
/db\_xref="taxon:9986"  
/cissue\_type="ileum"  
/dev\_stage="adult"  
1..41  
/partial  
42..1085  
/codon\_start=1  
/product="ileal sodium-dependent bile acid transporter"  
/protein\_id="CAA9184.1"  
/db\_xref="GI:1019395"  
/db\_xref="SWISS-PROT:Q28727"  
/translation="MSNLTVGCLANATVEGASCVAPBSFNALISVLSTVLTLLA  
LWFSWGNVIEIKKIFGHRIRPWGIFIGLCOFGIMPLGFLYLAVAFGIMPOAVVL  
IMGCGGTASNLINLAWVDGMDLSVMTGCTSLIAGMPLCLVYTTMTWDSGTV  
IYDNTGTSVLALVYVSVISGMPNPKKPKAKITLKVSGIAGAVLILAVVGGILYQ  
SMATPEKMTITGITTPMAYSIGFLIARIAGOPWTRKCTVALETGMONTQICSTTVQ

3'UTR  
BASE COUNT 265 a 274 c 283 g 294 t  
ORIGIN  
LSFSPEDITVFTPLPLYSIFQIAFPAIFGIYVANKKGNDAEPDIDKOTRE  
SSFHOMGGFQPE"  
1086..1116  
/partial

Query Match 27.0%; Score 306.4; DB 4; Length 1116;  
Best Local Similarity 59.2%; Pred. No. 1.6e-82;  
Matches 523; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

Qy 80 ATGGAACCTGGAGTCTGTTTTCACAGTGTGTCCACTGTGATGATGAGGGGCTGCATGT 139  
Db 124 ATGCCATCTCAGCGCTGTGTTCTGAGTACCGGTGACCATCTGTGCTGCTGCTGCTG 183  
Qy 140 TCTCTTGGAGATGTTCCGTGAGATCCGGAAGCTGTGTCACATCAGAGAACCTTGG 199  
Db 184 TCTCCATGGATGCACTGGAATCAAGAAATTCCTGGGCACTAAGCGGCGCTTGG 243  
Qy 200 GCATTCGTGGGAGCTGCTGTCAGATTTGGGCTTCATGCTTTTACAGCTTATCTCTGG 259  
Db 244 GCATCTTATGCTGCTTCTGTCAGATTTGGGATCATGCTCCCTCAGGGAATTTGCTTAC 303  
Qy 260 CCATTAAGCTTTTCTGTGAAGCCAGTCCAGCTATTTGCTGTCTATCATATGGGCTGCTCC 319  
Db 304 CGTGGCCCTTTGGAGATCCATCCATCCAGCGCGGTGTGTCTCATATGAGGATGCTCC 363  
Qy 320 CGGGGGGACCATCTCTAATATTTTCACTTCTGGGTTGATGAGATGATGATCTCAGCA 379  
Db 364 CTGGAGGAACGGCTCCCAACATCTCGGCTTATGGGTGATGAGACATGACCTTGATG 423  
Qy 380 TCAATGACCAACCTGTTTCCACCGTGGCGCTGGGAATGATGCACTTGCATTTATC 439  
Db 424 TCAGCATGACCACTGCTCCACATTTGCTGCTGCTGCAATGATGCTGATGCTTTATG 483  
Qy 440 TCTAAGCTGTGCTCGAGTCTTACAGAGAAATCCACATCTTATGAGAAATGAGAA 499  
Db 484 TCTACACCAAAATGAGGTGAGACTCTGGAGCCATGTAATCTTACAGATACATAGGTA 543  
Qy 500 TTACCTTGTGCTGACCAATCTCTGTGGCTTGTGTGTGATGATGATGATGATGATG 559  
Db 544 CTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603  
Qy 560 CAAACCAATCCAAATCATTTCTCAAGATTGGGGCGGTTTGGGGGCTCTCTCTGCG 619  
Db 604 CCAAAAGCGCAAGATTATTAAGTTGATGATGATGATGATGATGATGATGATGATG 663  
Qy 620 TGGTGGCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679  
Db 664 TCATGCTGCTGATGAGAGAAATATTTGACCAAGGTGCTGATGATGATGATGATG 723  
Qy 680 TGAACCATGATTTCACTTTCTTTGATGCGCATGTCACGGGTTTCTGCTGGCACTT 739  
Db 724 GGAATTATAGGAAGATATTTCCATGGCGGTATCTCCCTTGGCTTTTGGCTAGGA 783  
Qy 740 TTACCCACAGCTTGGCAAGGTGCAAGCAATTTCTTAGAACTGAGCTGAGATA 799  
Db 784 TAGCTGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 843  
Qy 800 TTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 859  
Db 844 CACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903  
Qy 860 TGAATTTCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 919  
Db 904 TCACCTTCCCGCTCATTTTACAGCATCTTCCAGATGCGCTTGGCACAATCTCTTAGGA 963  
Qy 920 CATATCAGAGTACAGAGAGATTTGAAGAACAAATGAGAA 963  
Db 964 TATATGTCATATAGGAATGTCATGAGAAATAATGATGAGAA 1007

RESULT 10

AR033871 AR033871 1047 bp DNA linear PAT 29-SEP-1999  
LOCUS Sequence 3 from patent US 5869265.  
DEFINITION AR033871  
ACCESSION AR033871  
VERSION AR033871.1 GI:5949476  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1047)  
AUTHORS Dawson, P.A.  
TITLE Ideal bile acid transporter compositions and methods  
JOURNAL Patent: US 5869265-A 3 09-FEB-1999;  
FEATURES Location/Qualifiers  
source 1..1047  
/organism="unknown"  
BASE COUNT 251 a 251 c 255 g 290 t  
ORIGIN  
Query Match 26.3%; Score 297.8; DB 6; Length 1047;  
Best Local Similarity 58.5%; Pred. No. 7.2e-80;  
Matches 518; Conservative 0; Mismatches 367; Indels 0; Gaps 0;  
QY 80 ATGAAACCTGAGCTCGTTTACAGTGTGCACTGTGATGATGGGCTGCTCATGT 139  
DB 80 ATAACTCTTAAGTGTGCTTAAGTACGATGATGATGATGATGATGATGATGAT 139  
QY 140 TCTCTTTGGGATGTTCCGTGAGATCCGGAAGCTGTGTGCGACATCAGAGACCTGGG 199  
DB 140 TCTCATTGGGATGCAACGTGGAATCTAAGAAATTTCTAAGGCACTAAGACCGGCTGGG 199  
QY 200 GCATTGCTGTGGGAGCTGCTGCGAGTTGGGGCTCATGCTTTTACAGCTTATCTCTGG 259  
DB 200 GATTGTGTGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259  
QY 260 CCATTAGCTTTTCTCTGAAGCCAGTCCAAAGCTATTGCTGCTCATCATGAGGCTGCTGCC 319  
DB 260 CGGTGGCTTTTGAACATCTCCGCTCCAGGCGGTAGTGTGCTCATTAAGATGTGCTCC 319  
QY 320 CGGGGGGACCATCTCTAATTTTCACTTCTGCGGTATGAGATATGATGATCTCAGCA 379  
DB 320 CTGGAGGAAGCTGCTTCATATCTTGGGCTATGAGTGTGATGCGACATGAGACCTGAGG 379  
QY 380 TAGATATGCAACCTGTTCCACGTCGAGCGCCCTGGGAATGATGCACTCTGATTTATC 439  
DB 380 TAGATATGCAACCTGTTCCACGTCGAGCGCCCTGGGAATGATGCACTCTGATTTATC 439  
QY 440 TCTACACCTGTCTGTGAGTCTTTCAGACAGATCTCAACATCTCTTATCAGAACATAGAA 499  
DB 440 TCTATACCAAAATGTGGTGTGAGCTGTGGAGCATGTATTCCTATATATACATAGATA 499  
QY 500 TTACCTGTGTGCTGACCATTTCTGTGCGCTTGTGTGTCTATGTGAATTCAGATGGC 559  
DB 500 CATCTCTGTGTGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 559  
QY 560 CAAAACATCCAAATCATTTCAAGATTGGGGCGGTGTGTGTGGGGCTCCCTCTTGG 619  
DB 560 CCAAAAGCAAGATCATCTTAATTAATGTGGTCAATGCGGCGCATCTCATGTGTGC 619  
QY 620 TGTGCGAGT 679  
DB 620 TATATGCTGT 679  
QY 680 TGACCATCATGTTTCACTTTCTTGTGATTTGCGCATGTCACGGCTTTTCTGTGCGCATTT 739  
DB 680 GATTATATAGAACATATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 739  
QY 740 TTACCAACAGTCTGTGGCAAGGTGAGGACAAATTTCTTAGAAATCTGAGCTCAGATA 799  
DB 740 TTGCTGT 799  
QY 800 TTCAAGTGTGATCAGCATGCTCACTTATCTTTCACTGTGAGGACCTGTGTGCTCAGATGT 859

DB 800 CGACATATGTTCCACATCTGTTAGCTCTCTCTTCACTCTGAGGAGCTCAATGTCTAT 859  
QY 860 TGAGTTTCCCATCTGGCCATATGATCTCTTCCAGCTATAGATGATTTCTTATGTGTGAG 919  
DB 860 TCACCTTCCGCTCATCTACAGCATTTTCCAGCTGCTTGTGCGCAATATTTCTTAGAT 919  
QY 920 CATATCAGACGTACAAAGAGATTTGAAGAACAAATCATGAGAAAA 964  
DB 920 TTTATGTGATACAAAGAAATGTCTATGAGAAAAACAAGCAGAAA 964  
RESULT 11  
LOCUS 132745 1047 bp DNA linear PAT 06-FEB-1997  
DEFINITION Sequence 3 from patent US 5589358.  
ACCESSION 132745  
VERSION 132745.1 GI:1823536  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1047)  
AUTHORS Dawson, P.A.  
TITLE Ideal bile acid transporter compositions and methods  
JOURNAL Patent: US 5589358-A 3 31-DEC-1996;  
FEATURES Location/Qualifiers  
source 1..1047  
/organism="unknown"  
BASE COUNT 251 a 251 c 255 g 290 t  
ORIGIN  
Query Match 26.3%; Score 297.8; DB 6; Length 1047;  
Best Local Similarity 58.5%; Pred. No. 7.2e-80;  
Matches 518; Conservative 0; Mismatches 367; Indels 0; Gaps 0;  
QY 80 ATGAAACCTGAGCTCGTTTACAGTGTGCACTGTGATGATGGGCTGCTCATGT 139  
DB 80 ATAACTCTTAAGTGTGCTTAAGTACGATGATGATGATGATGATGATGATGAT 139  
QY 140 TCTCTTTGGGATGTTCCGTGAGATCCGGAAGCTGTGTGCGACATCAGAGACCTGGG 199  
DB 140 TCTCATTGGGATGCAACGTGGAATCTAAGAAATTTCTAAGGCACTAAGACCGGCTGGG 199  
QY 200 GCATTGCTGTGGGAGCTGCTGCGAGTTGGGGCTCATGCTTTTACAGCTTATCTCTGG 259  
DB 200 GATTGTGTGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259  
QY 260 CCATTAGCTTTTCTCTGAAGCCAGTCCAAAGCTATTGCTGCTCATCATGAGGCTGCTGCC 319  
DB 260 CGGTGGCTTTTGAACATCTCCGCTCCAGGCGGTAGTGTGCTCATTAAGATGTGCTCC 319  
QY 320 CGGGGGGACCATCTCTAATTTTCACTTCTGCGGTATGAGATATGATGATCTCAGCA 379  
DB 320 CTGGAGGAAGCTGCTTCATATCTTGGGCTATGAGTGTGATGCGACATGAGACCTGAGG 379  
QY 380 TAGATATGCAACCTGTTCCACGTCGAGCGCCCTGGGAATGATGCACTCTGATTTATC 439  
DB 380 TAGATATGCAACCTGTTCCACGTCGAGCGCCCTGGGAATGATGCACTCTGATTTATC 439  
QY 440 TCTACACCTGTCTGTGAGTCTTTCAGACAGATCTCAACATCTCTTATCAGAACATAGAA 499  
DB 440 TCTATACCAAAATGTGGTGTGAGCTGTGGAGCATGTATTCCTATATATACATAGATA 499  
QY 500 TTACCTGTGTGCTGACCATTTCTGTGCGCTTGTGTGTCTATGTGAATTCAGATGGC 559  
DB 500 CATCTCTGTGTGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 559  
QY 560 CAAAACATCCAAATCATTTCAAGATTGGGGCGGTGTGTGTGGGGCTCCCTCTTGG 619  
DB 560 CCAAAAGCAAGATCATCTTAATTAATGTGGTCAATGCGGCGCATCTCATGTGTGC 619  
QY 620 TGTGCGAGT 679

```
Db      620 TCATAGCTGTGGTGGAGAAATATGTACCAAGCGCTGGATCATTTGCTCCCAACTGT 679
      680 TGACCATGAGTTTCACTTTTCCCTTTGATTTGGCCATGTCAGGGATTTTGTGTGGCACTTT 739
      680 GATATATAGGAACAATATTTCTGTGGCGGTTACCTCGGGGTTTCTTGTGGCAGAA 739
      740 TTACCAACCAAGTCTTGGCAAGGTGAGACAAATTTCTTAGAAGTGAAGCTCAGATA 799
      740 TTGCTGTGTACCTGTGTACAGGTGCGCAACGGTCTCTTTTGAACGGGAGATGCAAGAA 799
      800 TTCAGATGTCATCAGCATGCTCAGTTATCTTCACTGTGAGCACTTGTCAGATGT 859
      800 CCGAGTATGTTCCACATGCTTACCTCTTCTTCACTGAGAGCTCAATGTCTAT 859
      860 TGAGTTTCCCATGCTGCTTATGACCTCTTCCAGCTGATAGATTTCTTATTTGTCAG 919
      860 TCACCTTCCCGCTCATCTACAGCATTTTTCAGCTGCTTGGCCGAAATATTTAGAT 919
      920 CATATGACGCTGCAAGAGAGATTGAAGAACAAATGAGAAA 964
      920 TTATGTGGCATACAGAAATGTATGAGAAAACAGGACAGAA 964

RESULT 12
LOCUS   HSU10417              3779 bp      mRNA      linear      PRI 02-OCT-2001
DEFINITION Homo sapiens ileal sodium-dependent bile acid transporter
            (SLC10-A2) mRNA, complete cds.
ACCESSION U10417
VERSION   U10417.1
KEYWORDS  GI:2623285
SOURCE    Homo sapiens
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3779)
  AUTHORS Wong,M.H., Oelkers,P. and Dawson,P.A.
  TITLE Identification of a mutation in the ileal sodium-dependent bile
  JOURNAL acid transporter gene that abolishes transport activity
  MEDLINE J. Biol. Chem. 270 (45), 27228-27234 (1995)
  PUBMED  7592981
REFERENCE 2 (bases 1 to 3779)
  AUTHORS Craddock,A.L., Love,M.W., Daniel,R.W., Kirby,L.C., Walters,H.C.,
  TITLE Expression and transport properties of the human ileal and renal
  JOURNAL sodium-dependent bile acid transporter
  MEDLINE Am. J. Physiol. 274 (1 Pt 1), G157-G169 (1998)
  PUBMED  98120406
REFERENCE 3 (bases 1 to 3779)
  AUTHORS Dawson,P.A.
  TITLE Direct Submission
  JOURNAL Submitted (07-JUN-1994) Paul A. Dawson, Internal
            Medicine/Gastroenterology, Bowman Gray School of Medicine, Wake
            Forest University, Medical Center Boulevard, Winston-Salem, NC
            27157, USA
            4 (bases 1 to 3779)
            Dawson,P.A.
            Direct Submission
            Submitted (17-NOV-1997) Paul A. Dawson, Internal
            Medicine/Gastroenterology, Bowman Gray School of Medicine, Wake
            Forest University, Medical Center Boulevard, Winston-Salem, NC
            27157, USA
REMARK   Sequence update by submitter
COMMENT  On Nov 17, 1997 this sequence version replaced gi:595398.
FEATURES             Location/Qualifiers
     source            1..3779
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13q33"
                     /clone="PHISBT"
```



Db 1338 TTGCTGCTTACCTGCTAGACGTCGACGAGTGTCTTTTGAACGGGAGTGCAGAAC 1397  
Qy 800 TTGAGTGTGATACCATGCTCCAGTATCTTTTCACTGCTGAGCATTTGATCCAGATG 859  
Db 1398 CGCGCTATGTTCCACCATGCTGCTCTCTTCACTGAGGAGCTCATATGCTAT 1457  
Qy 860 TGAATTTCCCACTGGCTATGAGTCTTTCAGCTGATGAGGATTTCTTATTTGCGAG 919  
Db 1458 TCACCTTCCGCTCATCTACAGCATTTTCCAGCTGCTGCGCATATTTAGGAT 1517  
Qy 920 CATATCAGACGTACAGAGAGATTGAAGAAACAAACATGAGAAAA 964  
Db 1518 TTTATGTGCTACACAGAAATGTCATGAGAAAAACAGGACAGAAA 1562

RESULT 13  
LOCUS RNU07183 4269 bp mRNA linear ROD 16-JUL-2001  
INITIATION Rattus norvegicus ileal sodium-dependent bile acid transporter  
mRNA, complete cds.  
U07183  
ACCESSION U07183.2 GI:14719390  
VERSION  
KEYWORDS Rattus norvegicus.  
SOURCE Rattus norvegicus.  
ORGANISM Rattus norvegicus.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 4269)  
AUTHORS Shneider, B.L., Dawson, P.A., Christie, D.M., Hardikar, W., Wong, M.H.  
and Suchy, F.J.  
TITLE Cloning and molecular characterization of the ontogeny of a rat  
ileal sodium-dependent bile acid transporter  
J. Clin. Invest. 95 (2), 745-754 (1995)

JOURNAL MEDLINE 95164708  
PUBMED 7860756  
REFERENCE 2 (bases 1 to 4269)  
AUTHORS Christie, D.M., Dawson, P.A., Thevananthar, S. and Shneider, B.L.  
TITLE Comparative analysis of the ontogeny of a sodium-dependent bile  
acid transporter in rat kidney and ileum  
Am. J. Physiol. 271 (2 Pt 1), G377-G385 (1996)

JOURNAL MEDLINE 96365877  
PUBMED 8770054  
REFERENCE 3 (bases 1 to 4269)  
AUTHORS Chen, P., Ma, L. and Shneider, B.L.  
TITLE Transcriptional regulation of the rat apical sodium-dependent bile  
acid transporter: Basic mechanisms and response to inflammation and  
corticosteroids  
Unpublished

JOURNAL REFERENCE 4 (bases 1 to 4269)  
AUTHORS Shneider, B.L.  
TITLE Direct Submission  
JOURNAL Submitted (25-FEB-1994) Benjamin L. Shneider, Pediatrics, Yale  
University, 333 Cedar Street, New Haven, CT 06520, USA  
5 (bases 1 to 4269)  
Shneider, B.L.  
REFERENCE Direct Submission  
JOURNAL Submitted (11-JUL-2000) Pediatrics, Mount Sinai School of Medicine,  
One Gustave L. Levy Place (Box 1656), NY, NY 10029, USA  
REMARK Sequence update by submitter  
COMMENT On Jul 16, 2001 this sequence version replaced gi:1488241.  
FEATURES  
location/Qualifiers  
1..4269  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="BS37C1"  
/tissue\_type="ileum"  
/dev\_stage="adult"  
1..107  
108..1154  
5'UTR  
CDS  
/product="ileal sodium-dependent bile acid transporter"

/protein\_id="AAC53101.1"  
/db\_xref="GI:677882"  
/translation="MDSNVSPNAPFCEGDSCLVSNENALISTVMSVLTITLILAM  
VMSMGNVNEINKFLEHIIKRPKFIPIFQFLQRFIMPLTGTIILSVASILLVQAVYLI  
MGCGPGSGSNILAVIWDGMDLSVSTKSTSTILALMPLCLPIYIRKMWDSSTLYI  
PYDSIGSLVALIPIVSIIMFVNHKPKKIIKIGSISGAILIIVLAVAGGILYS  
AMIIPELMIIGITPIPIAGSLGFLARIGOPMYRRTALFEGMONTOLCSITVL  
SESPEDLNVTEPEPLIYTFVQLVFAIILIMVYTYKKCHGKIDAEFLEKIDNDMP  
SPEETNKGFPQDEK"  
BASE COUNT 1315 a 780 c 850 g 1324 t  
ORIGIN  
Query Match 25.9%; Score 293.2; DB 10; Length 4269;  
Best Local Similarity 59.1%; Pred. No. 2.3e-78;  
Matches 502; Conservative 0; Mismatches 348; Indels 0; Gaps 0;  
3'UTR  
Qy 109 GTGTCACTGTGATGAGGAGCTGCTCATGTTCTTTGGAGATGTCGAGATCCG 168  
Db 216 GTGCTCAACATTTCTTAGCCATGATGTTTCTATGGGGTGCATATGTAATCAAC 275  
Qy 169 AAGCTGTGTGCAATCAGAGAACCTGGGGCATTTGCTGTGGACTGCTTGCAGTTT 228  
Db 276 AAGTTCCTAGGACACATPAAAGCGGCATGCGGCACTTCTGTGGCTTCTCTGCACTT 335  
Qy 229 GGGCTCAAGCCTTTTACAGCTTATCTCGGCAATGACTTTTCTGGAAGCCAGTCCA 288  
Db 336 GGAATCATGCTCTCAGAGATTTATCTGTCTGTGGCTTGGCACTCTTCTGTGAG 395  
Qy 289 GCTATGCTGTTTTCATCATGAGGCTGCTGCCGGGGGACCATCTCTAACATTTTCAAC 348  
Db 396 GCTGTGGAGGTTAATATGATGAGTGTGCTGCCGGAAGAACCTGCTCAATATCTGAGC 455  
Qy 349 TTCTGGGTGATGAGATATGATATCTCAGATCATGATGATGATGATGATGATGATG 408  
Db 456 TATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 515-  
Qy 409 GCCCGGGAATGAGGACCTGTGATTTATCTTACACCTGCTCGAGTCTTCAAGCAG 468  
Db 516 GCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 575  
Qy 469 AATCTACCATTTCTTATGACATGATGATGATGATGATGATGATGATGATGATGATG 528  
Db 576 ACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 635  
Qy 529 GCTTGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 588  
Db 636 TCCATGGAATGTTGTAATACAAATGCCCCCAAGGAAAGATTAATTAATTAAT 695  
Qy 589 GGGGCGGTGTTGGTGGGCTCTCTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGG 648  
Db 696 GGAATCATGCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 755  
Qy 649 AAGAGATTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708  
Db 756 CAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 815  
Qy 709 GGCATGTACAGGTTTCTGCTGCACTTTTACCAACAGTCTTGGCAAGGTGAGG 768  
Db 816 GGCATACAGCTTGTGTTCTTCTGCTGATGATGATGATGATGATGATGATGATGATG 875  
Qy 769 ACAATTTCTTGAAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATG 828  
Db 876 AAGATGCTTGAAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATG 935  
Qy 829 TCTTCACTGTGAGCATTTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 888  
Db 936 TCTTTAGCCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 995  
Qy 889 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 948  
Db 996 CAGCTGCTTGTGAGCATATATATGATGATGATGATGATGATGATGATGATGATGATG 1055

QY	949	AACAAACATG	958
Db	1056	AAAAATGATG	1065

RESULT 14  
AC120684

LOCUS	106811 bp	DNA	linear	HTG 23-JUL-2002
DEFINITION	Rattus norvegicus clone CH230-1M5,	***	SEQUENCING	IN PROGRESS ***
REMARKS	52 unordered pieces.			

ACCESSION	AC120684
VERSION	AC120684.2
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Rattus norvegicus.

REFERENCE  
AUTHORS

1 (bases 1 to 106811)

Murny D.M., Adams,C., Adu-Oduola, B., Ali-osman, F.R., Allen C., Albrooks, S.L., Amatrungne, H.C., Are, J.R., Ayele, M., Banke, T., Barbarta, J., Benton, J., Blinage, K., Blankenburg, K., Bonini, D., Bowck, J., Bowles, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Butnah, C., Butch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chaver, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denna, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durkin, J.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gattisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gotsell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hayes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Lounseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Madeshwari, M., Mawpa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabhat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogun, M., Okunnu, G., Oregunye, N., Oviedo, R., Pace, A., Payton, B., Peery, U., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoochari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telifido, B., Thomas, N., Thomas, S., Usmani, K., Vaequez, L., Vera, V., Villalton, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, S., Williamson, A., Wlarczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

REFERENCE  
AUTHORS

2 (bases 1 to 106811)

Worley, K.C.

REFERENCE  
AUTHORS

3 (bases 1 to 106811)

Worley, K.C.

REFERENCE  
AUTHORS

Submitted (23-Jul-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Submitted (23-Jul-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

**COMMENT**

On Jul 18, 2002 this sequence version replaced gi:20514546

----- genome center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 52 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of 'N', but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved

1	1673:	contig of 1673 bp in length
1674	1773:	gap of unknown length
1774	3396:	contig of 1623 bp in length
3397	3496:	gap of unknown length
3497	4886:	contig of 1350 bp in length
4847	4966:	gap of unknown length
4647	6033:	contig of 1087 bp in length
6034	7330:	gap of unknown length
6134	7730:	contig of 1597 bp in length
7731	7830:	gap of unknown length
7831	9114:	contig of 1284 bp in length
9115	9214:	gap of unknown length
9215	10568:	contig of 1354 bp in length
10569	10668:	gap of unknown length
10669	11701:	contig of 1033 bp in length
11701	11801:	gap of unknown length
11802	13342:	contig of 1541 bp in length
13343	13442:	gap of unknown length
13343	14661:	contig of 1219 bp in length
14662	14761:	gap of unknown length
14762	16156:	contig of 1335 bp in length
16157	16256:	gap of unknown length
16257	17487:	contig of 1211 bp in length
17488	17567:	gap of unknown length
17568	18876:	contig of 1309 bp in length
18877	18976:	gap of unknown length
18977	21088:	contig of 2112 bp in length
21089	21188:	gap of unknown length
21189	22267:	contig of 1079 bp in length
22268	22367:	gap of unknown length
22368	23405:	contig of 1038 bp in length
23406	23505:	gap of unknown length
23506	24559:	contig of 1054 bp in length
24560	24659:	gap of unknown length
24660	26033:	contig of 1374 bp in length
26034	26133:	gap of unknown length
26134	27715:	contig of 1582 bp in length
27716	27815:	gap of unknown length
27816	29020:	contig of 1205 bp in length
29021	29120:	gap of unknown length
29121	30180:	contig of 1060 bp in length
30181	30280:	gap of unknown length
30281	31631:	contig of 1351 bp in length
31632	31771:	gap of unknown length
31732	33324:	contig of 1593 bp in length
33325	33424:	gap of unknown length





```

/chromosome="5"
/clone="RP23-514"
/clone_lib="RPct-23"
BASE COUNT      59453 a 48517 c 47922 g 59318 t
ORIGIN

```

```
/clone="RP23-514"
```

48517 C 47922 G

BASE COUNT	59453	a	48517	c	47922	g	59318	e
ORIGIN								

## ORIGIN

Query Match	23.1%	Score 261.8;	DB 10;	Length 215210;
Best Local Similarity	80.9%;	Pred. No. 2e-68;		
Matches 305;	Conservative 0;	Mismatches 72;	Indels 0;	Gaps 0;

Best Local Similarity 80.9%; Pred. No. 2e-68;

Matches 305; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Oy	1	ATGAGAGCCAAATGGTTCTTCAGACAGCTAGCCTGCGCCAGCAACAGTTCCAGAGGAGAGCTG	60
Db	139053	ATGAGACACAGACTGTGTCGGGGCACTCCACTGCTCCCTGTCAACAGTACGAGAGGAAGACCG	139112
Oy	61	CCAGTGGGACGTGAGGTGCATGGAAAACCTGAGAGCTGGTTTCA CAGTGGTGTCACACTGTG	120
Db	139113	CCCGTGGGAATGAGAGGGCCATGGAGATCTTAAAGCTCCTTTTACAGTGTCTCTGGCTGTG	139172
Oy	121	ATGATGGGGCTGCTCATGTTCTCTTTGGGATGTTCCGTGGAGATCCGGAAGCTGTGCTG	180
Db	139173	ATGGTGGGTTTGATCATGTTCTCTTTTGGATGTTCTGTGGAGAGCTAGAAAGCTCTGGTTG	139232
Oy	181	CACATCAGGAGACCCCTGGGGGCAATGCTGTGGGACGTCTCGACGATTTGGGCTCATGCC	240
Db	139223	CACCTCAGAAAGACCTCTGGGGACATCGATGAGGCTGCTCTTCCAGTTTGAATTATAGCCT	139292
Oy	241	TTTACAGCTTATCTCTCGGCACATTAAGCTTTTCTGTGAAGCAGTCCAAAGTATGTCTGTT	300
Db	139293	CTGACAGCTTATCTGTATTACCATTTGGCTTCGGTCTGAACACANTTCCAAAGTATTTGCTGTG	139352
Oy	301	CTCATCATGCGCTGCTGCCCGGGGGGACCAATCTTAAACATTTTCACTTTCTTGGGTTGAT	360
Db	139353	CTCATGATGGGGAGCTGCCCTCTGGGGGACCAATCTTAAAGTTCTTCACTTCTGGGTTGAT	139412
Oy	361	GGAGATATGATCTCAG	377
Db	139413	GGAGATATGATCTCAG	139429

Db 139053 ATGAGCACAGACTGTGCGGCACTCCACCTGCCCTGTCAACAGTACGGAGGAGACCCG 139112

61 CCAGTGGGACTGGAGGTCATGGAACCTGGAGCTCGTTTTCACAGTGGTGTCCACTGTG 120

Db 139113 CCGTGGGATGGAGGGCCATGCCAATCTAAAGCTGCTTTTACAGTGTCTCGGCTGTG 139172

121 ATGATGGGCTGCTCATGTTCTTTGGAGATCCGAGAGCTGTGTCG 180

139173 ATGGTGGGTTTGATCATGTTCTTTTGATGTTCTGTGGAGAGTCAGAAGCTCTGGTTG 139232

181 CACATCAGGAGACCTGGGCAATGCTGGGACTGCTCTGCCAGTTGGGTCATGGCT 240

Db 139233 CACCTCAGAAGACCCCTGGGCACTCCAGTGGGCTTCCTTCCAGTTGGACTTAAGCCT 139292

241 11ACAGC1A1C1C1GGCA1A5C111C1C1GAAGCCAG1CCAGC1A115C1911 300

DD 139293 CIGACAGC IATCIG IAGCCAI IGGC ICGG CIGAAACCAI ICCAAGCAI IGCIGC 139352

2000

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

27 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258

DB - 139413 GGAGATATGGATCTCAG 139429

Search completed: June 9, 2003, 06:27:26  
Job time : 3078 secs

Job time : 3078 secs